U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Admin.
National Ocean Service
Office of Response and Restoration
c/o EPA Office of Site Remediation and Restoration (HIO)
1 Congress Street
Boston, MA 02114
26 September 2003

Ms. Susan Alison Wolfe, MNG Center at SRA
2801 Clarendon Blvd., Suite 100
Arlington, VA 22201

Thank-you for the Ecological Risk Assessment for GE/Housatonic River Site, Rest of River, (July, 2003) prepared by Weston. NOAA was very impressed with the number of assessment endpoints and the substantial well-designed measurements. NOAA looked for specific issues that are of concern to our agency mission . Hence, this review considers:
A. Appendix F: Assessment Endpoint - Survival, Growth, and Reproduction of Fish.
B. Section 3: Assessment Endpoint - Community Structure, Survival, Growth, and Reproduction of Benthic Organisms and Appendix D – Benthic Invertebrates.
C. Section 5: Assessment Endpoint - Survival, Growth, and Reproduction of Fish.

A. Appendix F: Assessment Endpoint - Survival, Growth, and Reproduction of Fish.

1. F.2.2.2, p. F-16. Congeners and TEFs used to calculate TEQ should be provided in a table. Assuming the WHO (van den Berg et al. 1998) TEFs were used, were all 29 PCB, dioxin and dibenzofuran congeners measured? If not, what is the potential for underestimation?

2. F.3.2.1, p. F-27. Note the statement: "Therefore, the effects thresholds below 10 mg/kg tPCB were not considered appropriate for application to warmwater fish in the PSA." This appears to be a somewhat arbitrary decision. The decision is based on "relative sensitivity" to PCBs, yet the references cited to support this assumption refer to dioxin-like effects on fish. The authors present no evidence to support the assumption that reproductive and developmental effects from PCBs are all related to AhR-mediated toxicity. In addition, the species they eliminate as "too sensitive" are also used in studies that were accepted. Perhaps it would be more appropriate to evaluate the exposure and endpoints used in the tests in terms of sensitivity. For example, exposing eggs to PCBs 9 days pre-hatch (Mauck et al 1978) is not equivalent to long-term maternal exposure (e.g., Hendricks et al 1981), which is a more realistic and appropriate laboratory exposure. Although the authors have classified each study as affecting "reproduction" or "development," the studies within these classifications have different degrees of relevance for deriving the "effects threshold."

3. Attachment F.4, p. 3-9: Limiting the literature compilation to studies with whole body PCB measurements greatly reduces the number of studies available, particularly those

with most relevant exposure and toxicity endpoints. The authors in the "Data Gaps and Uncertainties" section acknowledged this: "There were a limited number of studies that met the screening criteria used in the review of the literature." And: "studies replicated the maternal transfer of contaminants and related adult tissue concentrations to effects in offspring. In studies that did address maternal transfer, eggs or fry concentrations were frequently reported, rather than adult whole body." In a weight-of-evidence type analysis, the relevance of the exposure and the sensitivity of the endpoint should be primary weighting factors before developing a threshold effect level.

The authors cite the range in Monosson (1999) of 25-70 mg/kg Aroclor 1254 in the liver to support their selected threshold (31 mg/kg), but the reproductive endpoints addressed are very different. The studies included by Monosson refer to the "functioning of the hypothalamic-pituitary-gonadal-liver axis (HPGL)," compared to juvenile survival and growth. Monosson (1999b) also states: "In addition, several studies have shown that embryos and larvae are even more sensitive to A1254 than adult fish. It should be noted that the endpoint used in embryo or larval studies is often reduced survival. This suggests that there may be sublethal effects at even lower concentrations than those found to cause embryo or larval toxicity (such as growth or behavior). Laboratory studies have shown embryo or larval survival to be affected by concentrations as low as 5 ppm."

NOAA believes that field studies may represent a more realistic exposure than laboratory studies and should be considered more directly in this evaluation. Our reasoning is that PCBs in the environment may be considerably different from original commercial mixtures used in the laboratory due to weathering in the environment and differential accumulation in the food web. Also, field studies represent long-term exposure, whereas lab exposure to post-hatch larvae is typically to PCB-free water.
Field studies where PCBs are a major contaminant (e.g., New Bedford Harbor, Sheboygan River) should be considered.

Reference:

Monosson, E. 1999b. Reproductive and developmental effects of PCBs in fish: a synthesis of laboratory and field studies. Reviews in Toxicol. 3: 25-75.

B. Section 3: Assessment Endpoint - Community Structure, Survival, Growth, and Reproduction of Benthic Organisms and **Appendix D** – Benthic Invertebrates.

1. Reflecting the comments above from Section F4, the report contains several number of references to data, but no clear citation to where the reader can review those. An example is on page 3-17, where in the box at the bottom of the page the data sources are listed, but without citation. In addition, the report presents the results of substantial data manipulations and only in a few cases refers the reader to a description of the data used and details of the procedures employed.

2. Appendix D (Assessment Endpoint – Benthic Invertebrates) seemed much more readable than the main report, but still suffers from a lack of depth and clarity in

presenting key details of the data used. The report seems to stop at a number of points where more complete consideration of the results and supporting data might substantially enhance the determination of risk.

3. p. D-62: "Contaminant Mixtures – SQVs are most functional and predictive when used to evaluate the potential adverse effects of single contaminants. However, a complex mixture of contaminants is present at many contaminated sites. The confounding effects of the various contaminants can affect the ability of the SQVs to accurately predict toxicity for a single contaminant." This unsubstantiated statement of the authors' opinion only applies to theoretical (e.g., equilibrium partitioning) Sediment Quality Guidelines (SQGs). Empirically-derived SQGs, which represent the majority of the SQGs under consideration, implicitly account for contaminant mixtures. Additionally, the analyses presented for the most part ignore contaminants other than PCBs.

We could not find any overall evaluation of the reference or site stations that summarized all contaminants compared to SQGs (e.g., the number of contaminants exceeding benchmarks such as ERM or the maximum ERM quotient) to get an overall assessment of the degree of contamination from contaminants other than PCBs at all locations.

4. A more specific concern about the lack of an overall perspective in the report is that there appears to be substantial portions of the results that are not supportive of each other. The major factors we identified were 1) the differences in the PCB concentrations and spatial trends in concentration with river mile among the laboratory/triad studies, the *in situ* toxicity studies, and the general river assessment data; and 2) toxicity was observed at stations without altered benthic communities, while altered benthic communities were observed at stations with no or limited toxicity. The report did not adequately address the somewhat dramatic differences in the data, nor justify how these data can be used together**.** The same PCB concentrations are associated with one kind of effect at one location, but not the other effect, and vice-versa. The report appears to ignore that this lack of concordance.

NOAA's interpretation is that the authors are making the argument that the PCB concentrations collected for the toxicity tests and with the benthic community samples are site-associated, but unique to those samples. Thus, the data derived from each portion can be compared to the PCB concentration without regard to the wider body of data that might be associated with that location, including samples collected at another time from the "same location." On that basis, the results can then be used to develop exposure-responses that in turn lead to threshold concentrations and MATCs that can then be used with the wider body of data. The underlying lack of similarity in PCB concentrations, however leads to some question as to the representativeness of all of the data.

5. The derivation of the 3 mg PCB/kg from the toxicity testing data is perhaps a good example of limits in the reports presentation. The authors state they selected the lowest six LC50s/EC50s and averaged those (we could not find a lowest six in the tables of the main text of Appendix D that gave us the averages presented). The reason why six were chosen is not stated. Using the approach a number of different six-test averages were

developed depending on which reference sediment was used for comparison. In addition, the same procedure was used to generate similar LC20-based averages, as well as averages using the "synoptic PCB concentrations" rather than the location average. All these potential MATCs were compared and the value of 3 mg PCB/kg "was selected" (page 3-49.lines 26-31). As described, the process seemed subjective. Neither the reasonableness of the choices, nor the viability of other choices was presented. For example, the authors note that the longer-running tests tended to give lower LC50s/EC50s. It seems reasonable to select those results preferentially as more representative of the benthic exposure, which is certainly long-term. Why not consider the endpoints available and select the most relevant?

6. Similarly, the analysis of the benthic community data is performed almost completely using only combination of metrics, in the multivariate analyses, or metrics that are themselves "combination" metrics, i.e., total abundances and taxon richness. No detailed examination of the underlying data are presented to demonstrate the reasons for the differences (or lack there of) among locations, not the reasonableness of concluding that the changes are likely to associated with PCB expose. Even univariate analysis of the six metrics used in the mulitvariate analysis would be helpful as a means of better identifying what characteristic of the benthos is driving the differences. As the report notes, there are many uncontrolled and potentially unidentified habitat-related and other factors that could cause similar differences.

7. It was not clear why the MHBI analysis was included at all. The authors discounted the results, as was appropriate for this index of generic organic matter enrichment.

8. The measurement endpoints state (page 3-8, line 4) that "species-specific indications of adverse effects" are included. No such evaluation is presented.

9. A number of conceptual figures of the study design and information flow could be clarified. As a quick example, Figure 3.1-1, page 3-3, has a triple line and no arrow connecting the contaminated river media to "Food Sources." This designation is not included in the legend. In addition, the "Media" line includes biota tissue as a source of direct uptake by the benthic community.

10. Another one of those detail things. In the box at line 19, page 3-14, mention is made that the median concentrations were used for data collected at the same location on the same day, to avoid bias in temporal variability. Yet temporal variability is not discussed. Spatial variability is only alluded to, but never discussed quantitatively except as variance over broad reaches of the river. There was very little attempt to quantify the effects of variability in sediment concentrations on the results of the studies. This concern is exemplified by the discussion in Section 3.2.4.2.2, page 3-18, which notes that data collected from the same location but at different times, in this case the PCB concentrations measure with the benthic community samples and the PCB concentrations measured at the same stations for the toxicity testing, "should" not be used for other than the study for which it was collected. This restriction would seem to bring into question

the use of any of the data collected in the river as representative of other than a "snap-shot" of conditions in a small location at on instance in time.

11. To make the spatial comparisons of the PCB concentrations in the sediments simpler, it would be very helpful if the river mile locations were included with the location Ids for the benthic community and the toxicity testing samples (Figures 3.2-3, -4, and –5, page 3-19—21).

12. Given the interest of the ERA to benthic organisms, it seems inappropriate to include the Daphnia magna test results. This organism is not benthic. Similarly, it also seems inappropriate to include the *Lubriculus variegates* toxicity endpoints when the authors state that this is organism is know to be tolerant. Toxicity results from neither test appeared to have influenced the conclusions of this section, but their inclusion only adds unnecessary complexity to the report.

C. Section 5: Assessment Endpoint - Survival, Growth, and Reproduction of Fish.

1. The overall feel was similar to Section 5 as the authors were presenting overviews of a substantial body of data and data manipulations that were not cited or adequately described. As was the case with the benthic ERA, we felt that there were additional details in data and data interpretation that could have been explored to make the fish ERA more precise and compelling. Specific types of information that might be helpful include more detailed presentation and discussion of the spatial distributions of PCB concentrations, including congeners, in fish species in the PSA and reference areas; more detailed evaluation of the pathologies observed in the toxicity tests; and more consideration of the age structure and recruitment to the fishers in the PSA indicated in the field studies.

Overall, as was the case in Section 3, NOAA is impressed that the report presents risk information from a number of independent approaches that concur on the presence and approximate level of risk to fish in the ERA.

2. The conceptual model identifies water-born exposure as a route to the fish, but the measurement endpoints do not include any comparison to water-exposure TRVs. This exclusion seems odd and limiting, especially since the benthic ERA at least considered this route of exposure. In general it is accepted that dietary exposure is more important, and the ERA focuses very clearly on fish tissue concentrations as the basis for determining risk. For completeness and comparability, however, a simple consideration of water exposure is appropriate.

3. We had the same problem with the Figures 5.1-1 to 5.1-3 that we did with the comparable figures in the benthic ERA—the flow of information and the descriptions in the boxes do not seem complete. The simple box of measurement endpoints on page 5-8, line 18+ is much easier to understand.

4. It would be helpful to match the studies listed on page 5-9 to the measurement endpoints they support, as well as provide citations to the data/studies listed.

5. The report should state specifically in Section 5.2.2, page 5-11, what data were used, including at least a citation. As written, this section implies but does not specifically state that the non-EPA data were excluded. Appendix F states that both EPA and GE data were used, but the source(s) are not cited. The rationale for exclusion seems weak, i.e., that the additional data would not change the results. In most instances, more data provides a more robust conclusion.

6. The data presented indicate that the concentrations of PCBs in whole-body-reconstituted samples were higher than the PCB concentrations in whole-body samples (Tables 5.2-1 and 5.2-2, and Figure 5.2-1). The report does not indicate whether the difference is considered real (statistically significant) and whether the difference is the result of bias from the analytical approach or because of size/age differences in the fish treated by the different approaches. Because the species have data of both sample types (as well as composite samples), it would be useful to know more about how these data were used in comparing to the MATCs developed later.

7. As was the case with the data reduction in the benthic ERA, the fish ERA appears to use the reasoning that combining all results to grand average yields a "better" result. This approach is evident in the analysis of the literature data, where a threshold concentration was selected (page 5-25, line 18) that is the average concentration associated with effects in the studies compiled. On the face of it this approach may be more robust, but it begs the question as to whether one or more of the studies/endpoints of the studies are more appropriate or relevant. However, it is likely that the Weight of Evidence approach (not reviewed in this report by NOAA) will correct for this concern.

8. Similarly, the pathology data from the toxicity studies were treated primarily as "total abnormalities," and not differentiated or examined in the report to determine whether one or more specific "abnormality" is most strongly associated with a PCB effect, is truly pathological, or correlates better with the concentration data. It appears that different fish showed different responses that were pooled/added to develop and overall rate of effected fish. It was not clear to us that this treatment is valid without more explanation of why one would not expect some similarities in responses by the fish. This concern is supported by the paragraph on page 5-34, line 12, which states the some "deformities" were "not reliable markers for PCB exposure." Were those abnormalities included in the totals?

9. It appears that a different approach to interpreting the data was used to develop the effects threshold for TEQ (page 5-28, line 1), than was used for the tPCBs. No rationale for this difference is given.

10. In the fish toxicity studies, the number of exceptions and failed tests alluded to provide a high level of confidence in those results. An example is the data presented in Figure 5.3-3, showing the control substance triolein having as great an effect as most of

the doses, and the non-linear dose-responses observed in most of the other similar figures. In addition, the results presented in Figures 5.3-8 and –9 include notations for results where the "dose response was unusual." Variability is expected, but it needs to be discussed clearly. In this report, I found it very hard to determine the degree of consistency among the study results.

11. It was not clear why a 50% response level was selected for deriving threshold concentrations from the fish studies, rather than a more protective level.

12. NOAA agrees with the overall conclusion regarding the fish population studies and related studies, that they demonstrate that the effects of exposure in the PSA are not "catastrophic." However, as noted above it would have been more satisfying if the authors could have provided more discussion of the age, size, trophic structure of the fish community, together with GE's data on LM bass nesting and spawning success, to refine the risk to sustainable normal populations in the PSA. For example, the nesting surveys seem to imply that the hatching success was not good (page 5-52, line 26).

13. As noted above, the data interpretation/analyses did converge on similar threshold concentrations (page 5-54, box at line 16). These results do provide substantial support for an effect concentration in the range demonstrated. The data presentation (Figures 5.4-1 to –6 and Tables 5.4-1 and –2) comparing these threshold concentrations to the concentrations in fish tissue was nicely done and provided a good illustration of the extent/likelihood of risk to the different species.

14. The uncertainty analysis provided in Section 5.4.6 suffers from the same limitations noted for the benthic ERA: the possible effect of the uncertainty on the risk statements is not discussed. Using the first bullet on page 5-71, line 11, as the example, is there no information regarding what kinds of seasonal changes would be expected to predict whether the results presented earlier likely over- or under-estimate the PCB exposure?

15. Section 5.4.7.1 fails to note that an evaluation of the risks to fish in Reach 8 was included earlier in the chapter (Section 5.4.3.1.2, page 5-55).

Please let me know if you need further information.

Sincerely,

Kenneth Finkelstein, Ph.D.

CC: Bryan Olson (EPA)
Susan Svirsky (EPA)
Ken Munney (USF&WS)